Clutch Performance in Sport and Exercise: A Systematic Review

Date Submitted: 1st November, 2019

Date Resubmitted: 4th May, 2020

Abstract

Improved performance under pressure in sport and exercise has been termed clutch performance. The aim of this study was to systematically review, synthesise, and evaluate existing research on clutch performance. Specifically, this review explored: (i) research designs used to examine clutch performance; (ii) definitions of clutch performance; (iii) theoretical frameworks underlying clutch performance; (iv) how clutch performance has been measured; (v) the level of supporting evidence for clutch performance; and, (vi) evidence regarding how clutch performances occur. Ten electronic databases were searched in October 2019, with 27 studies found to meet the eligibility criteria included in the review. The results indicate that there is considerable definitional, conceptual, and measurement heterogeneity in the field of clutch performance. Multiple, conflicting definitions of clutch performance were identified in the literature, which consequently led to the adoption of two distinct approaches to examining clutch performance as: (i) an ability; or, (ii) an isolated episode of performance. These differing approaches have resulted in disparate measurement strategies, and accordingly, there was mixed evidence for the concept of clutch performance and how it occurs. In response to these issues, we propose four principles to help guide future research towards refined explanations of clutch performance.

Keywords; performance under pressure; subjective experience; psychology; optimal performance; anxiety

Clutch Performance in Sport and Exercise: A Systematic Review

Increased performance under pressure in sport and exercise has been referred to as *clutch performance* (Otten, 2009; Swann et al., 2019). The term clutch performance is frequently applied by the media to many high-profile, celebrated sporting moments, such as Michael Jordan scoring with five seconds remaining to win the 1998 National Basketball Association (NBA) Championship (Woodyard, 2018); the New England Patriots' 31-point, second half comeback to win the 2017 Super Bowl (Hurley, 2019); and Sergio Aguero's injury time goal to win Manchester City's first Premier League title in 2012 (Hart, 2017). Recent evidence suggests that such clutch performances are intrinsically rewarding and motivating (Swann et al., 2017a), and that clutch performances can also occur in exercise settings (Swann et al., 2019). As these performances occur under pressure, clutch performance has been considered psychological in origin (Otten, 2013). Facilitating clutch performance is therefore of great interest to researchers and practitioners in the field of sport and exercise psychology (Marchant et al., 2014; Otten, 2013)

The phrase 'in the clutch' was first used in a 1929 New York Times article to describe when a baseball batter hits a safe 'blow' at an opportune moment (Safire, 2005). Despite having a long history of colloquial use (e.g., West & Libby, 1969), scientific definitions of clutch performance have only emerged relatively recently. The most prominent definitions of clutch performance are those provided by Otten (2009) and Hibbs (2010). Otten (2009) defined clutch performance as 'any performance increment or superior performance that occurs under pressure circumstances' (p. 584). Hibbs (2010), meanwhile, defined clutch performance as:

> when a participant in competitive sport succeeds at a competitive-related, challenging task during a clutch situation, is aware that the performance occurs during a clutch situation, possesses the capacity to experience clutch situation-related stress, cares

> about the outcome of the contest, and succeeds primarily due to skill rather than luck or cheating (p. 55).

A clutch situation, according to Hibbs (2010), is ‘a point in a competitive sport where the success or failure of the participants has a significant impact on the outcome of the contest’ (p. 48). Researchers have highlighted, however, that definitions of clutch performance remain problematic. For example, Seifreid and Papatheodorou (2010) noted that ‘clutch exists as a challenging concept which is inadequately defined in sport’ (p. 92), whilst Mesagno and Hill (2013) stated that clutch performance is ‘ambiguously defined’ (p. 275). Swann et al. (2017a), meanwhile, suggested that ‘standard definitions of clutch performance may require refinement’ (p. 2278). Definitional critiques have also centered on the situations in which clutch performances occur, based on evidence that clutch performances have been reported outside of competitive sport settings, such as training (Swann et al., 2017a) and in exercise contexts (Swann et al., 2019). As such, questions remain over how to adequately define clutch performance, as well as the situations in which such performances occur.

Theoretical explanations of clutch performance have emerged from two different approaches. Traditionally, theories of performance under pressure have focused on choking, defined as ‘an acute and considerable decrease in skill execution and performance when self-expected standards are normally achievable, which is the result of increased anxiety under perceived pressure’ (Mesagno & Hill, 2013, p. 274). For example, attentional theories propose that, in response to anxiety, athletes either divert attention towards the self (e.g., self-focus theories; Beilock & Carr, 2001), or away from task-relevant cues (e.g., distraction theories; Oudejans et al., 2011). More recently, an Integrated Model of Flow and Clutch States has been proposed (Swann et al., 2017b, 2019). This model outlines that a specific psychological state may underlie clutch performance (i.e., clutch states), which overlaps with, yet is distinct from, the experience of flow (a deeply focused, absorbing, and autotelic

experience; Csikszentmihalyi, 2002). As such, explanations of clutch performance have emerged out of research centred on either choking or flow.

A range of measurement approaches have been adopted to examine clutch performance. Research in this field began with Cramer's (1977) investigation into the existence of clutch hitters in baseball. For the subsequent 30 years, clutch performance research was exclusively conducted within the sport of baseball, through the method of sabermetrics (i.e., the statistical analysis of baseball; Costa et al., 2019). Generally, such archival approaches have typically focused on whether clutch performance exists as an observable phenomenon in sport. In the last decade, however, there has been a considerable increase in the quantity and diversity of research examining clutch performance. For example, measurement approaches have extended to include qualitative methodologies that focus on the psychological state underlying clutch performance (e.g., Swann et al., 2017a), whilst experimental approaches have included measuring variables such as subjective experience (e.g., anxiety), technique changes in sport-specific skills (e.g., golf-putting stroke), and objective performance (e.g., putting accuracy) during clutch performances (e.g., Gray & Cañal-Bruland, 2015). In parallel, research has expanded into a wide range of sports beyond baseball, such as basketball (e.g., Otten, 2009), golf (e.g., Hill & Hemmings, 2015), and tennis (e.g., Jetter & Walker, 2015), as well as exercise (e.g., Swann et al., 2019).

There are fundamental questions surrounding the strength of evidence underpinning clutch performance as an observable phenomenon in sport. For example, Wallace et al. (2013) found no evidence for NBA players displaying clutch performances during the fourth quarter of playoff games. Similarly, Birnbaum (2008) demonstrated that clutch performance in Major League Baseball (MLB) was not a predictor of future clutch performances, casting doubt on the notion that certain players are more prone to producing clutch performances than others. In contrast, Jetter and Walker (2015) found that higher-ranked professional

tennis players improved their winning percentage, both overall and in decisive sets (i.e., tiebreak sets), during important competitions (i.e., Grand Slam tournaments). This finding suggested that higher-ranked players are able to produce clutch performances when the incentives were greatest. Meanwhile, Solomonov et al. (2015) indicated that NBA players with reputations for being clutch players (i.e., known for producing repeated clutch performances) increased their output (e.g., points scored) in the last five minutes of critical games. However, these players' overall base performance (e.g., shooting percentage) did not increase. Solomonov et al. (2015) concluded that this finding provided limited evidence of clutch players, in that whilst these players scored more points, this was a consequence of shooting more often, rather than improved shooting accuracy. Thus, there is contradictory evidence as to whether clutch performance exists in sport.

Against the backdrop of definitional issues and conflicting evidence, a systematic review of clutch performance is both timely and important in terms of providing guidance on future directions for the field. Systematic reviews aim to be 'comprehensive, methodical, explicit, transparent, and as unbiased as possible in the questions they explore and how they explore them' (Siddaway et al., 2019, p. 97). Thus, systematic reviews aim to produce a summary of the literature that explores relations, contradictions, and gaps in a research field and the reasons for these. In turn, systematic reviews can allow broad and more robust conclusions to be drawn, which can outline future research directions and inform practice (Siddaway et al., 2019). Furthermore, systematic reviews have previously been employed as a method to review and bring clarity to constructs with definitional issues in the field of sport and exercise psychology (Dohme et al., 2017; Swann et al., 2015). These aspects are highly relevant to the field of clutch performance, which has yet to be systematically reviewed and synthesised, and may benefit from greater clarity and direction.

The aim of this study was to systematically review, synthesise, and evaluate the existing research on clutch performance. Specifically, this review addressed the following research questions: (i) what research designs have been used to examine clutch performance?; (ii) how has clutch performance been defined?; (iii) what theoretical frameworks have been used to explain clutch performance?; (iv) how has clutch performance been measured?; (v) is there supporting evidence for clutch performance in sport and exercise?; and, if so, (vi) what is known about the occurrence of clutch performances? In turn, this review seeks to address existing issues currently facing the field by providing definitional and conceptual clarity. Further, this review aimed to identify future directions for research on clutch performance, which can increase understanding of how practitioners, athletes, and exercisers can facilitate successful performance under pressure.

Methods

Protocol

The review was guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009). The PRISMA checklist is reported in Supplementary File 1. The search strategy included 10 electronic databases, representing a combination of sport- (SPORTDiscus) and psychology- (PSYCInfo, PSYCArticles) specific databases, and general scientific databases (Academic Search Complete, SCOPUS, Pub Med, Medline, Web of Science, Science Direct, ProQuest Central). The final search was conducted in October 2019.

Potential search terms were initially developed by the authors, all of whom have published in the area of clutch performance. Combinations of these search terms were trialed by the first author on the EBSCOhost database. These preliminary searches were reviewed for relevance, and the search repeated until the most effective combination of search terms were identified (Siddaway et al., 2019). The aim of this process was to limit the amount of

irrelevant results, whilst ensuring all relevant literature was retained. The final search string was: [clutch] AND [(sport* OR exerci* OR physical* OR athlet*)]. The singular use of the term *clutch*, rather than *clutch performance*, was chosen to capture terminology relevant to the concept, but that may not contain the term performance (e.g., hitting in the clutch, clutch shooting). The search terms physical* (e.g., physical fitness) and athlet* (e.g., athlete) were included as synonyms to supplement sport* and exerci*. Exercise was included in this review as recent evidence suggests that clutch performances may also occur in exercise settings (e.g., Swann et al., 2019). Where possible, the first block was searched in the title, abstract, and keyword field, whilst the second block was searched in the full text field. The full search strategy for each database is presented in Supplementary File 2.

Eligibility Criteria

Inclusion and exclusion criteria were employed to ensure that the scope of the review was clearly defined, and that all literature relevant to the aims of the review was identified (Siddaway et al., 2019; Centre for Reviews and Dissemination, 2009). Criteria for inclusion were that articles must: (a) be a peer-reviewed journal article published in the English language; (b) report original empirical evidence (including original analyses of secondary data); (c) be published prior to October 2019 (when the final search was undertaken); and, (d) examine the nature, existence and/or occurrence of clutch performance in participants' engaging in sport[1] (including sport-specific skills) or exercise[2], as defined by the World Health Organisation (WHO, 2018). Articles were excluded that (e) referred to clutch as a mechanical apparatus (e.g., a clutch in motorcycle sports). Following initial scoping of the

[1] 'An activity involving physical exertion, skill and/or hand-eye coordination as the primary focus of the activity, with elements of competition where rules and patterns of behaviour governing the activity exist formally through organizations; and may be participated in either individually or as a team' (WHO, 2018, p. 101)

[2] 'A subcategory of physical activity that is planned, structured, repetitive, and purposive, in the sense that the improvement or maintenance of one or more components of physical fitness is the objective' (WHO, 2018, p. 98)

196 literature, inclusion of original analyses of secondary data were deemed important for the
197 current review. Specifically, archival studies comprise a significant portion of the extant
198 literature, and consideration of these studies is pertinent to several aims of the review (e.g.,
199 how clutch performance has been measured).

200 **Screening Process**

201 Following database searching, articles were imported and screened in Endnote X8
202 reference management software (Thomas Reuters, California), during which duplicates were
203 automatically removed. Missed duplicates during this stage were removed manually during
204 the screening process. Articles were independently screened at the title, abstract, and
205 keyword level for relevance by the first and third author. Studies were retained if they
206 contained the term clutch in the title, abstract, or as a keyword, appeared to involve
207 participants in the domain of sport or exercise, and were not referring to clutch as a
208 mechanical apparatus (e.g., in motorsports). A number of steps were followed to ensure that
209 the screening process was as comprehensive as possible (Siddaway et al., 2019). If the
210 relevance of an article was uncertain, the full text was obtained for further screening. Once
211 full texts were obtained for all identified studies, a further manual search was conducted by
212 the first author. Specifically, reference lists of all identified studies were searched, in addition
213 to forward searching citations of identified studies using Google scholar. This process was
214 repeated with each new study added. Lastly, authors who had two or more first-author
215 publications at this stage of screening were contacted and asked to suggest any relevant
216 literature that was not presently included (Siddaway et al., 2019). This resulted in two
217 additional studies (Jackman et al., in press.; Maher et al., 2018) being included, which had
218 been published after the initial search date. After completing these steps, the first and third
219 authors screened the full texts in accordance with the eligibility criteria. In three cases
220 inclusion was uncertain (Cramer, 1977; Cramer & Palmer, 2008; Deane & Palmer, 2006)

because it was not initially clear if original data had been analysed. Upon repeated readings and discussions, the reviewers agreed to include these papers as it was determined that original data had been analysed.

Data Extraction and Synthesis

Data were extracted by the first author. These data included: (i) study characteristics (methodology, study design, aims, hypotheses, theoretical framework); (ii) participant characteristics (sample size, gender, mean age, sport, expertise); and (iii) key findings relevant to the aims of the review (definitions, existence and occurrence of clutch performance). Given the heterogenous nature of the included studies, a narrative synthesis was undertaken. A narrative synthesis summarises and explains findings textually (Popay et al., 2006), with the aim of generating new insights (Thomas et al., 2012). A preliminary synthesis was initially conducted by tabulating textual summaries of the data according to the review aims. Tabulation is valuable in developing initial summaries of the included studies, as well as facilitating identification of patterns across studies (Higgins et al., 2019). Following this preliminary synthesis, the relationships between studies were explored by examining factors that may explain differences in findings between studies (Popay et al., 2006). This was an important step as two of the five review aims related to empirical findings. An interpretative approach was taken, in which findings of the included studies were filtered according to the conceptual assumptions and methods adopted (Drisko, 2019). Specifically, this involved examining how research design, definitions, and measurement may have informed the results of individual studies.

Quality Appraisal

Study quality was appraised using the 16-item assessment tool (QATSDD) developed by Sirriyeh et al. (2012) [3]. The QATSDD can be used to assess the quality of qualitative, quantitative, and mixed methods studies. However, criterion 14 of the tool was excluded on grounds of being ineffective for assessing reliability in qualitative research (Jaarsma & Smith, 2018; Smith & McGannon, 2018), whilst criterion 9 of the tool was excluded when scoring archival studies, as this criterion was deemed inappropriate for archival designs by the research team.

To limit bias, and facilitate transparency and trustworthiness, authors of the present review who were also authors on an included study were not involved in the quality assessment of that study. As such, the first author assessed 26 of the 27 studies, whilst the second, third, and fourth authors all assessed eight studies each. For the remaining studies, two independent reviewers were used. The first independent reviewer assessed four studies (three in conjunction with the first author, one in conjunction with the second independent reviewer), whilst the second independent reviewer assessed one study. All studies were assessed by two reviewers. As outlined in Sirriyeh et al. (2012), the reviewers met to discuss and deliberate on any scoring differences, following which a final score was determined by mutual agreement.

Results

In total, 4779 studies were identified across three separate searches. Following duplicate removal, 2548 studies were independently screened for relevance. The majority of studies screened at this stage were removed as they were not in the domain of sport or exercise (clutch is a prominent term in the fields of zoology and mechanical engineering). This process left 34 studies to be screened at the full text stage. An additional manual search

[3] To ensure the most appropriate tool was selected, three appraisal tools were piloted with five of the included papers, which were of a diverse methodology. These were the QATSDD (Sirriyeh et al., 2012), Mixed Methods Appraisal Tool (Pluye et al., 2011), and the QualSyst (Kmet et al., 2004). Following piloting, the QATSDD (Sirriyeh et al., 2012) was considered the most appropriate tool for the present review.

identified 14 potentially relevant articles to be screened at the full text stage. Thus, 48 articles were screened at the full text stage. Following full text screening, 21 articles were excluded. Reasons for exclusion were that the studies: were not original empirical research (*n* = 11); did not examine the nature, existence and/or occurrence of clutch performance (*n* = 5); were not peer reviewed (n = 2); were not in the domain of sport or exercise (*n* = 2); and, were not written in English (*n* = 1). Accordingly, 27 articles were included in the systematic review. The PRISMA diagram of this process is provided in Figure 1.

INSERT FIGURE 1 AROUND HERE

Characteristics of Included Studies

Details of study characteristics, including type of sport/exercise, sample size, methodology, methods, approach to research design, and key findings relevant to aims of the review are presented in Table 1. In total, 17 studies were quantitative, six qualitative, and four mixed methods. Of the quantitative studies, 13 employed archival methods, whilst the remaining four studies used experimental methods. In the qualitative studies, both career-based and event-focused[4] semi-structured interview methods were used. Three mixed method studies used a combination of psychometric measures and interviews (see Table 1 for measures), whilst one mixed methods study (Swann et al., 2016) included performance observation, naturalistic performance data, and event-focused, semi-structured interviews.

INSERT TABLE 1 AROUND HERE

There were 545 (304 male, 241 female) participants from studies that collected primary data. Data were observed for at least 3652[5] individuals from studies that obtained secondary data (i.e., archival methods). Meanwhile, six studies did not report the sample size

[4] Career-based interviews seek general understanding of a phenomenon over an athlete's career or significant period of time (Swann et al., 2018). Event-focused interviews collect data soon after one specific event (e.g., within hours/days), which allows for more detailed and chronological recall of the event (Swann et al., 2018)

[5] The sample size from Otten & Barrett (2013) was not included in this calculation, as it was unclear how many athletes appeared more than once (e.g., as pitching, batting, and team statistics were calculated for multiple seasons, meaning the same athlete may have been observed more than once)

in adequate detail to report. Participants were examined in a range of sports, including: baseball (n = 8); basketball (n = 6); golf (n = 5); mixed sport (n = 3); tennis (n = 1); and American football (n = 1). A mix of participants engaging in both sport and exercise was examined in two studies (Swann et al., 2017a, 2017b), whilst only participants in exercise were examined by Swann et al. (2019).

Quality Appraisal

Table 1 also displays quality appraisal scores from the QATSDD (Sirriyeh et al., 2012) for the included studies. The mean quality appraisal score across all studies was 61%. Archival studies generally received the lowest quality scores, on account of lacking clear conceptual frameworks, not justifying sample sizes, and omitting discussion of strengths and weaknesses (a full score for each paper by category is found in Supplementary File 3). Experimental studies, meanwhile, ranged from scores of 50% (McEwan et al., 2012) to 71% (Otten, 2009). Qualitative and mixed method studies were generally the highest scoring and, with the exception of Owens et al. (2016; 38%) and Maher et al. (2018; 56%), all scored above 80% (see Table 1).

Research Design

There were two distinct approaches to how research was designed to examine clutch performance. The most common approach (n = 14) was to examine clutch performance over a series of related performances. For example, studies measured clutch performance across multiple games (e.g., Solomonov et al., 2015), consecutive seasons (e.g., Birnbaum, 2008), or entire careers (e.g., Deane & Palmer, 2006). These were primarily archival studies, but also involved one mixed methods study (Owens et al., 2016; see Table 1). Hibbs (2010) has previously termed this approach ‘clutch ability… when one is notable for delivering clutch performances’ (p. 48). Accordingly, we term this the *clutch ability* approach.

The other approach (n = 13) was to examine clutch performance in isolated episodes of performance. For example, studies investigated a single experimental session (e.g., Otten, 2009), an isolated performance (e.g., Swann et al., 2016), or a number of isolated performances, which were considered unrelated, from the same athlete (e.g., Jackman et al., 2017). Studies examining isolated performance episodes were experimental, qualitative, or mixed methods in design (see Table 1). We term this the *clutch episodes* approach. These two approaches represent different conceptual perspectives on clutch performance, and consequently, have implications for how it should be measured. As such, the remainder of this Results section will consider, where possible, these two approaches separately.

Defining Clutch Performance

Definitions of clutch performance from the included studies are provided in Table 2. An explicit definition of clutch performance (or related concepts, see *clutch ability, clutch situations, and clutch states*) was not provided in 26% (n = 7) of the studies. Clutch was defined in terms of a performance (i.e., a performance under pressure; Swann et al., 2017a), as an ability (i.e., the ability to produce repeated clutch performances; Deane & Palmer, 2006), a situation (i.e., a high pressure or critical game situation; McEwan et al., 2012), or a psychological state (i.e., the subjective experience underlying clutch performance; Swann et al., 2019). These different definitions are discussed below.

INSERT TABLE 2 AROUND HERE

Clutch performance. The most common definition (n = 10) of clutch performance was Otten's (2009) definition. This definition was the first instance in the included literature that clutch was defined in terms of performance, rather than in terms of an ability or situation. It is unclear, however, whether Otten's (2009) definition strictly refers to a singular performance episode. For example, two studies (Otten & Barrett, 2013; Solomonov et al., 2015), which measured clutch performance over multiple performances, employed Otten's

(2009) definition. Six studies referenced Hibbs' (2010) definition of clutch performance. Of note, five of these studies also referenced Otten's (2009) definition. In these five studies, both definitions were viewed as complementary (i.e., used together – see Table 2), rather than compared or contrasted. Indeed, none of the included studies examined the implication of using different definitions of clutch performance on the same data (i.e., if using different definitions changed the findings). Lastly, Maher et al. (2018) defined clutch performance as "adaptive (e.g., clutch) responses" (p. 1) to pressure. The definition employed by Maher et al. (2018) is considerably vague, and it is unclear how, or if, this definition fits with either Otten's (2009) or Hibbs' (2010) definition of clutch performance.

Clutch ability, clutch situations, and clutch states. Clutch was defined as an ability in four studies. Two of these definitions were specific to baseball (Cramer & Palmer, 2008; Deane & Palmer, 2006), with the remaining definitions generalisable across sports (Jetter & Walker, 2015; Owens et al., 2016 – see Table 2). Interestingly, Owens et al. (2016) cited Otten's (2009) definition, but clearly positioned clutch as an ability (i.e., 'a clutch athlete exhibits superior performance under pressure'; Owens et al., 2016, p.4). As above, it is unclear whether Otten's (2009) definition is episodic or can apply to studies examining clutch ability.

A definition of a clutch situation was provided in four studies. Baseball-specific definitions were provided in three of these studies (Birnbaum, 2008; Brooks, 1989; Ruane, 2005), whilst one study provided the broad definition of a clutch situation as 'instances of high pressure' (McEwan et al., 2012, p. 144). Clutch states, meanwhile, were defined as the psychological state underlying clutch performances (Jackman et al., 2017; Swann et al., 2017b). Whilst both Jackman et al. (2017) and Swann et al. (2017b) also provided definitions of clutch performance, it is unclear if clutch states and clutch performance are two distinct

constructs, or if they are interconnected (i.e., if the experience of clutch states is an inherent aspect of clutch performance, and vice versa).

Comment. To date, various approaches to examining and defining clutch performance have been employed in the literature. It is therefore important that consistent terminology is used for the remainder of the Results. Accordingly, *clutch performance* will be used as an umbrella term, incorporating both clutch ability (i.e., clutch performance over a series of related performances) and clutch episodes (i.e., clutch performance as an isolated performance episode). Where possible, the more specific terminology of either clutch ability or clutch episodes will be used.

Theoretical Frameworks and Clutch Performance

Table 1 provides an overview of the theoretical frameworks and conceptual models within the included studies. From the included studies, 33% (n = 9) provided no explicit theoretical framework for clutch performance. These studies may therefore be considered atheoretical. The following section discusses the different theoretical frameworks that were employed in the remaining studies.

Choking-based explanations. Eleven studies examined clutch performance in relation to choking. Primarily, these studies drew on attentional theories (n = 8), which included self-focus theories (n = 5), distraction theories (n = 1), or both self-focus and distraction theories (n = 2). Of note, the majority (n =5) of studies utilising attentional theories employed definitions that called for increased performance (e.g., Otten, 2009). No explanation was provided, however, for how such theories accounted for increased performance (i.e., only the proposed mechanisms behind performance breakdown were described). One study (Worthy et al., 2009), meanwhile, drew on regulatory focus theory. This theory explains that athletes are more likely experience performance decrements when trying to avoid losing the game, as opposed to trying to win the game. Lastly, Hill and

Hemmings (2015) and Hill et al. (2017) examined the self-presentation model. The self-presentation model is concerned with understanding how one's self-presentation motives affect their performance anxiety, which may then precede attentional breakdowns via self-focus or distraction.

Integrated Model of Flow and Clutch States. Six studies (Jackman et al., 2017, in press; Swann et al., 2016, 2017a, 2017b, 2019) positioned clutch states within the Integrated Model of Flow and Clutch States. This model outlines the performance contexts, process of occurrence (discussed further below), subjective experience, and outcomes of clutch states. Hence, this model focuses on explaining the psychological state underlying clutch performance, rather than clutch performance per se (see *Inadequate Theoretical Framework* for further discussion).

Neoclassic economic theory. One study (Cao et al., 2011) stated that 'neoclassical economic theory predicts that individuals exert the most effort, and consequently produce their best performances, when the returns to effort are highest' (p. 231). Little further information, however, was provided about this theory, and how the results may or may not support it.

Measurement of Clutch Performance

No established measure of clutch performance was utilised in the included studies. Accordingly, this section reviews approaches to measurement with respect to the two essential constructs of clutch performance (i.e., those constructs that are core across definitions of clutch performance): (i) performance; and, (ii) pressure.

Measuring performance. Naturally, performance is inherent in the study of clutch performance. The following section addresses approaches to measurement of performance in studies examining clutch ability, and in studies examining clutch episodes.

Clutch ability. Table 3 presents the ways in which performance was assessed in the included studies. Objective measures of performance were employed in the majority of studies assessing clutch ability (n = 13; 94%). These studies all examined archival, naturalistic performance data. The benchmarks against which performance was assessed ranged considerably, however, and included comparing performance against: career averages (Cao et al., 2011); previous season performance (e.g., Birnbaum, 2008); performance within the same season (e.g., Birnbaum, 2009); and, performance within the same game (e.g., Wallace et al., 2013). In one study performance was assessed against an athlete's projected performance (i.e., clutch ability was judged against performances that had not yet occurred; Deane & Palmer, 2006). Across all of these studies, performance was considered to have improved if there was a statistically significant increase compared to the respective performance benchmark (e.g., one's career average; Cao et al., 2011). Subjective measurement of performance, meanwhile, was adopted in one mixed methods study (Owens et al., 2016). In this study, performance was assessed by asking a coach to evaluate which players in their team consistently performed well, or did not perform well, under pressure.

INSERT TABLE 3 AROUND HERE

Clutch episodes. As displayed in Table 3, studies in which clutch performance was assessed as an isolated episode primarily measured performance using subjective methods (n = 8; 62%). Generally, measurement involved participant self-report through semi-structured interviews, which principally reported athletes' and exercisers' perceptions of their own performance.

All experimental studies utilised objective measures of performance (n = 4; 31%). Performance was assessed following pressure manipulation in a sport-specific task (i.e., golf-putting, n = 3; basketball free-throw shooting, n = 1), and then compared with baseline scores. In three studies (Gray et al., 2013; Otten, 2009; McEwan et al., 2012), performance

improvement following pressure manipulation was considered clutch performance. As in the archival designs, performance was considered to have improved if there was a statistically significant increase compared to baseline performance. One study (Gray & Cañal-Bruland, 2015) meanwhile, considered clutch performance to be evident in those participants who did not choke. Accordingly, the clutch performance group in this study still decreased performance relative to baseline, but to a significantly lesser degree than those who choked. This suggests confusion around the extent of the performance increment required to classify a clutch performance.

One study (Swann et al., 2016) utilised both objective and subjective measures of performance. Specifically, this study involved observations of professional golfers during the final rounds of tournaments, a performance monitoring tool to objectively 'indicate peaks and troughs in the player's performance' (p. 104), and then event-focused interviews about the same rounds as soon as possible afterwards. To date, this appears to be the only study that has combined both objective and subjective measurement of performance.

Measuring pressure. The construct of pressure is central to definitions of clutch performance. Pressure is 'the presence of situational incentives for optimal, maximal, or superior performance' (Baumeister & Showers, 1986, p. 362), and importantly, involves a subjective component. The following sections review approaches to measurement of pressure in studies investigating clutch ability, and studies investigating clutch episodes.

Clutch ability. Table 4 provides an overview of the methods used to measure pressure in the included studies. The majority of studies (n = 13; 94%) designed to measure clutch ability did not directly measure pressure. Instead, as a proxy measure, certain in-game situations were used to represent pressure. Across these 13 studies, eight different situations were specified to infer pressure (see Table 4). Generally, these were situations considered important to the overall outcome of the game or tournament, although there was some

460 inconsistency. For example, Solomonov et al. (2015) considered pressure in the NBA as the
461 last five minutes in games within a score differential of 6-points, in the last 20 games of the
462 regular season. Worthy et al. (2009), meanwhile, considered pressure as the last minute in
463 games within a score differential of 5-points, in NBA playoff games. Taken together, the
464 decisions to determine what situations and factors represent pressure seem rather inconsistent
465 and arbitrary. Indeed, only one study (Otten & Barrett, 2013) provided supporting
466 justification that the assessed situation – MLB playoff games – was likely to increase an
467 athletes' pressure. Specifically, Otten and Barrett (2013) noted that greater fan attendance,
468 media attention, and internal and external rewards were likely to increase traditional forms of
469 pressure (e.g., presence of audience, ego relevance, reward contingency; Baumeister &
470 Showers, 1986).

471 The remaining study that examined clutch ability utilised a mixed-methods design.
472 Owens et al. (2016) conducted a single coach interview, which involved the coach identifying
473 which players performed well under pressure. In addition, Owens et al. (2016) also
474 distributed a ProScan Survey (Professional Dynamic Programs, 2003) to athletes, who were
475 instructed to reflect on how they expect to perform under pressure. The ProScan Survey has
476 been validated as a measure of personality (Hubby & Williamson, 1988), though not as a
477 measure of pressure. In summary, it is difficult to discern the extent to which the
478 operationalisation and measurement of pressure was valid across studies examining clutch
479 ability.

480 INSERT TABLE 4 AROUND HERE

481 ***Clutch episodes.*** Studies designed to examine clutch episodes used a range of
482 methods and tools to measure pressure (see Table 4). Qualitative and mixed methods
483 approaches predominantly involved interviewing athletes and exercisers (n = 9). Interview

methods allow for rich and detailed descriptions of subjective experiences (Smith & Sparkes, 2019), and hence, could offer a valuable avenue for in-depth explorations of pressure.

Experimental studies (n = 4) primarily employed psychometric measures of anxiety to examine pressure. Gray et al. (2013) asked participants to respond to the Immediate Anxiety Measures Scale (IAMS; Thomas et al., 2002). Similarly, Gray and Cañal-Bruland (2015) used the cognitive and somatic anxiety items of the IAMS, which has been identified as a valid and reliable measure of anxiety (Thomas et al., 2002), whilst also assessing changes in participants' average heart rate between trials. Meanwhile, Otten (2009) employed the Competitive State Anxiety Inventory 2 (Revised) (CSAI-2R; Cox et al., 2003), which is also a validated measure of anxiety (Cox et al., 2003). Whilst anxiety has been identified as an indicator of pressure (e.g., Gucciardi & Dimmock, 2008), measures of anxiety do not directly measure perceptions of pressure (Kent et al., 2018). As such, it is arguably the case that these experimental studies did not actually measure pressure, but examined a single, negatively framed (e.g., Burton & Naylor, 1997), indicator of pressure. Lastly, McEwan et al. (2012) asked participants 'how much pressure and anxiety they felt throughout the experiment' (p. 145). Responses to this question, however, did not undergo formal qualitative analysis, and accordingly were not reported in the results. Hence, the validity of this pressure manipulation is unclear.

Evidence for Clutch Performance

Evidence for clutch performance as an observable phenomenon was mixed. This section reviews the evidence for clutch performance with respect to studies that examined clutch ability, and studies that examined clutch episodes.

Clutch ability. Ten studies explicitly investigated the existence of clutch ability in sport[6]. From these studies, eight did not provide support for the existence of clutch ability. In studies examining baseball, fluctuations in performance during pressure situations were demonstrated to be more likely a product of random variation (Brooks, 1989; Cramer & Palmer, 2008; Deane & Palmer, 2006; Ruane, 2005), general hitting quantity (Cramer, 1977), or in the case of pitching, other performance factors (e.g., run support; Birnbaum, 2009). Further, clutch performance in one season was not predictive of clutch performance in future seasons (Birnbaum, 2008). In basketball, meanwhile, Wallace et al. (2013) demonstrated that most players were statistically average during the 4th quarter of NBA playoff games when compared with the previous 3 quarters of the same game, indicating no evidence of clutch ability.

In contrast, Jetter and Walker (2015) demonstrated support for the existence of clutch ability in tennis. Higher ranked players were more likely to win a Grand Slam tournament relative to other events, and also more likely to perform well in clutch situations within the match (e.g., tie-breaks). Furthermore, Solomonov et al. (2015) showed that 'clutch players' performance generally improves in the sense that they exert more effort in the final, critical moments of the game' (p. 136). Metrics such as foul drawing, free throw attempts, and successful free throws significantly increased compared to earlier periods in the game. These findings raise questions as to what aspects of performance must increase to be considered a clutch performance. For example, is increased effort, or specific components of performance – such as fouls drawn – sufficient, or is a more global perspective of performance outcomes necessary for clutch performance? In summary, there was limited support for the existence of clutch performance when examined as an ability. The measurement limitations of these

[6] Not all studies designed to examine clutch ability explicitly investigated whether the concept existed. Rather, four studies (Otten & Barrett, 2013; Cao et al., 2011; Owens et al., 2016; Worthy et al., 2013) assumed *a priori* that clutch performance, or clutch ability, existed.

studies, however, should be considered when assessing the validity of this conclusion. Specifically, it is unclear to what extent pressure was experienced by athletes in these studies, and the performance benchmarks used to assess performance were inconsistent.

Clutch episodes. In contrast to studies examining clutch ability, studies investigating isolated clutch episodes demonstrated strong support for the existence of clutch performance. Experimental studies generally indicated that participants could increase performance in response to pressure manipulations (Gray et al., 2013; Otten, 2009; McEwan, 2012). Qualitative studies showed that athletes could recall having clutch performances (Hill et al., 2017; Hill & Hemmings, 2015; Maher et al., 2018), whilst at the experiential level, clutch states – the subjective experience of clutch performance – were reported to occur during excellent sport performances and rewarding exercise experiences (e.g., Jackman et al., 2017). Specifically, clutch states were proposed to consist of 12 characteristics: absence of negative thoughts; absorption; altered sensory perceptions; automaticity of skills; confidence; deliberate focus; enhanced motivation; enjoyment; heightened arousal; heightened awareness; intense effort; and perceived control (Jackman et al., 2017, in press; Swann et al., 2016, 2017a, 2017b, 2019). In summary, support for clutch performance both as a performance outcome, (e.g., Gray et al., 2013) and at an experiential level (e.g., Swann et al., 2017a), was demonstrated in studies examining clutch episodes.

Occurrence of Clutch Performance

This section reviews factors involved in the occurrence of clutch performances. Given that limited supporting evidence was found for clutch ability, this section focuses solely on the occurrence of clutch episodes.

Clutch episodes. From experimental studies, a range of factors were identified in the occurrence of clutch performance. Gray et al. (2013) demonstrated that in golf putting, participants who increased performance under pressure had improved putting kinematics (i.e.,

swing amplitude) compared to baseline performance. McEwan et al. (2012), meanwhile, showed that participants who warmed up under high-pressure conditions performed significantly better in a single-shot, golf-putting task than those who warmed up under low-pressure conditions. Lastly, Otten (2009) indicated that a sense of perceived control during a free-throw task was the strongest predictor of clutch performance. The factors identified in the occurrence of clutch performance, therefore, varied considerably across experimental designs, and included technique improvements, warm-up strategies, and psychological mechanisms.

The occurrence of clutch performance episodes was also investigated in qualitative designs. Hill and Hemmings (2015) reported a number of approach coping strategies to facilitate clutch performance, such as simulated practice, performance routines, and cognitive restructuring (e.g., re-appraising threatening stressors as a challenge). The roles of simulated practice and performance routines in the occurrence of clutch performance were also highlighted by Maher et al. (2018) and Hill et al. (2017), in addition to a range of other factors. For example, Hill et al. (2017) reported that a sense of perceived control and challenge appraisal were also involved in the occurrence of clutch performances. Collectively, factors that consistently emerged out of these qualitative studies were challenge appraisal, simulated practice, and performance routines.

Four studies reported the occurrence of clutch states as a series of steps (Swann et al., 2016, 2017b, 2019; Jackman et al., in press). Clutch states occurred in contexts characterised by importance, and when the participant was still in contention to achieve an important goal. Athletes and exercisers initially appraised the situation as a challenge before setting specific goals relating to the desired outcome of that situation. Athletes and exercisers then made a deliberate decision to ‘step up’ their effort and intensity in order to try and achieve those goals (Swann et al., 2019, p. 92). In addition, Jackman et al. (in press) reported that the

occurrence of clutch states occur may be related to an athlete's mental toughness. Specifically, athletes high in mental toughness reported a more rapid initiation of clutch states than athletes low in mental toughness, particularly when in response to setbacks. Whilst processes of occurrence for clutch states has been consistently reported (Swann et al., 2016, 2017b, 2019; Jackman et al., in press), questions remain over the relationship between clutch states and clutch performance (i.e., do clutch states always underlie clutch performances?).

Discussion and Recommendations

The aim of this review was to synthesise and evaluate existing research on clutch performance in sport and exercise. The findings indicated that research into clutch performance has gathered momentum in the last decade. Over 75% (n = 21) of the included studies were published since 2009, with a third (n = 9) published since 2016. This momentum suggests that clutch performance is a contemporary field of research in sport and exercise psychology (e.g., Perry, 2019). Findings from this review, however, also suggest there is significant definitional, conceptual, and measurement heterogeneity within the field. Clutch performance has been defined inconsistently, with definitions referring to this construct both as an ability and an individual performance, whilst studies have also employed definitions of clutch situations and clutch states. Accordingly, two major approaches are evident in the field, which conceptualise clutch performance as an: (i) ability; and (ii) individual performance episode. These differing approaches have resulted in disparate measurement of clutch performance with questionable validity, and consequently, conflicting evidence regarding the existence of clutch performance.

Assessing Evidence for Clutch Performance

Studies which explicitly investigated the existence of clutch ability (n =10) demonstrated limited support. As Hibbs (2010) noted, however, 'in order to assign clutch

ability to a competitor, one must first know what a clutch performance is' (p. 48). At present, definitions of clutch performance lack specificity and clarity (see *Definitional Issues*), and consequently, it is difficult to determine exactly what clutch ability is. Moreover, studies examining the existence of clutch ability relied on proxy measures of pressure (i.e., certain game situations were used to infer pressure), meaning that the extent to which these athletes experienced pressure is unclear. Against this backdrop of definitional and measurement issues, making any conclusions about the existence of clutch ability based on current literature seems somewhat premature.

In contrast, support for isolated episodes of clutch performance was demonstrated across qualitative, experimental, and mixed methods designs. These studies identified a variety of factors in the occurrence of clutch performance. For example, technique improvements (e.g., Gray et al., 2013), simulated practice and performance routines (e.g., Maher et al., 2018), and psychological processes (e.g., perceived control; Otten, 2009) were all identified in the occurrence of clutch performance. In addition, Swann et al. (2016, 2017b, 2019) and Jackman et al. (in press) highlighted a sequential process in the occurrence of clutch states. Whilst these studies provide evidence for isolated episodes of clutch performance, they also highlight inconsistencies in how the occurrence of clutch performance has been examined, ranging from exploration of distal factors (e.g., simulated practice; Maher et al., 2018) to more proximal factors (e.g., perceived control; Otten, 2009). This perhaps suggests that even within studies adopting a similar approach (i.e., clutch episodes), there remains some confusion over how to examine the occurrence of clutch performance.

Definitional Issues

Definitions are important in facilitating conceptual clarity, informing measurement, and determining the direction of future research (Cooper et al., 2001; Wacker, 2004). This review demonstrated that 26% (n = 7) of studies did not provide a definition of clutch, in any

sense. When definitions were provided, these extended beyond defining clutch performance, and were also provided in terms of an ability (i.e., the ability to produce repeated, increased performances during critical game situations; Deane & Palmer, 2006), a situation (i.e., performance situation which is high in pressure; McEwan et al., 2012), and as a psychological state (i.e., the subjective experience underlying clutch performance; Swann et al., 2019). These varied definitions suggest conceptual confusion surrounding what clutch performance is, and is not. The most common definitions of clutch performance, meanwhile, were applied inconsistently. Otten's (2009) definition of clutch performance was cited both in studies that examined clutch performance as an ability (e.g., Solomonov et al., 2015), and as an individual episode (e.g., Hill et al., 2017). Further, five studies supplemented Otten's (2009) definition with Hibbs' (2010) definition of clutch performance, despite there being meaningful differences between the two (see *Guiding Principles for Clutch Performance Research).* Hence, a key finding from this review is that current definitions of clutch performance have not facilitated conceptual clarity and, accordingly, may require refinement to clearly differentiate between clutch ability and clutch performance episodes.

Inadequate Theoretical Framework

Robust theory represents a fundamental aim of science, providing the foundation upon which research and practice should be built (Cunningham, 2013; Doherty, 2013). The present review indicated that current theoretical approaches to clutch performance are insufficient. The most popular approach (n = 11) within the included studies was to employ theories (i.e., attentional theories) and models (i.e., self-presentation model) that primarily focused on explaining the mechanisms underlying performance breakdown. Both attentional theories and the self-presentation model, however, are grounded in performance responses to anxiety. Whilst anxiety is an indicator of pressure (e.g., Gucciardi & Dimmock, 2008), it has not been demonstrated that experiencing pressure always results in anxiety. Indeed,

Baumeister and Shower's (1986) formative, and widely used (e.g., Low et al., 2020), definition of pressure is relatively neutral (i.e., 'the presence of situational incentives for optimal, maximal, or superior performance', p. 362). As such, it may not be the case that all clutch performances are preceded by symptoms of anxiety or occur in a state of anxiety. Therefore, based on current understandings of clutch performance, attentional theories and the self-presentation model do not account for the range of potential responses to pressure that may lead to clutch performance.

The Integrated Model of Flow and Clutch States (Swann et al., 2017b, 2019) was employed in six studies, and describes the occurrence and experience of clutch states. Whilst this model emerged from a primarily qualitative methodology based on inductive analysis, and is to undergo harsher tests (e.g., experimental designs), it does outline a process of occurrence for clutch states. Importantly, these predictions can be tested and, if unsupported, falsified. It remains unclear, however, if clutch states are inherent to clutch performance, and vice versa. As with attentional theories and the self-presentation model, the Integrated Model of Flow and Clutch States (Swann et al., 2017b, 2019) only provides a partial explanation of clutch performance (i.e., based on clutch states). Lastly, a third of the included studies (n = 9) employed no theoretical framework for clutch performance. This both limits the utility of these studies (i.e., cannot adequately explain and predict phenomena; Bacharach, 1989), and highlights that a notable quantity of clutch performance research has been atheoretical. In summary, current theories and conceptual models do not offer complete explanations of clutch performance. Future research, therefore, needs to work towards development of a specific theory of clutch performance.

Methodological Critique

Broadly defined constructs lacking in specificity and clarity may result in disparate measurement (Wacker, 2004). The impact of unclear definitions of clutch performance is

679 evident in the extent to which measurement has been approached inconsistently. Clutch
680 performance was examined as an ability in just over half of the included studies, which
681 primarily involved utilising archival designs. Measurement of performance in archival
682 designs ranged from comparing performance within the same game (e.g., Wallace et al.,
683 2013) to comparing performance with a career average (e.g., Cao et al., 2011), highlighting
684 the unclear nature of what benchmark clutch performance should be compared against.
685 Further, archival studies did not directly measure pressure. Instead, pressure was treated as a
686 categorical variable that was inferred from the performance situation (i.e., it was assumed all
687 athletes experienced the same amount of pressure in certain situations, such as all games
688 within a Grand Slam tournament; Jetter & Walker, 2015). Indeed, only one study (Otten &
689 Barrett, 2013) justified why the performance situation (i.e., MLB playoffs) inferred pressure.
690 This general lack of measuring pressure is problematic as pressure involves a subjective
691 component (Baumeister & Showers, 1986) and, therefore, it cannot be assumed that all
692 athletes will perceive these situations in the same way.
693 The impact of unclear definitions was also evident in experimental studies that
694 examined clutch performance as an isolated episode. For example, different performance
695 thresholds were used to categorise clutch performances between experiments (e.g., Gray &
696 Cañal-Bruland, 2015). This suggests a need for consensus over the performance level
697 required for clutch performance (i.e., increased or maintained performance). Furthermore, the
698 use of psychometric measures of anxiety to assess pressure is incomplete. Whilst
699 measurement of anxiety may indicate the intensity of cognitive and somatic anxiety, this
700 provides little information regarding how, or if, pressure is interpreted facilitatively. Indeed,
701 it is not clear whether the perception of pressure necessarily results in increased anxiety.
702 Accordingly, more complete measurement of pressure is important, especially when
703 considering questions have been raised about the capability of experimental designs to

replicate the demands of naturalistic pressure situations (Gucciardi & Dimmock, 2008; Jackson, 2013).

From the included studies, qualitative and mixed method approaches represent the most appropriate measure of pressure at present. This is because interviews allow an in-depth exploration of pressure following real-world episodes of clutch performance. These interview methods, however, differed in their methodological strength. Specifically, three studies (Hill & Hemmings, 2015; Hill et al., 2017; Maher et al., 2018) employed career-based interviews, which ask athletes to report on events that occurred months or years in the past (Swann et al., 2018). In contrast, event-focused interviews aim to interview athletes within hours or days of a performance and have been suggested as a methodologically stronger alternative (Swann et al., 2018). This is because event-focused interviews may reduce the risk of athletes' forgetting details or presenting a biased recall (Brewer et al., 1991; Yarrow et al., 1970). Accordingly, studies that employ single event-focused interviews (Jackman et al., in press; Swann et al., 2017b, 2017a, 2019, 2016) may offer the most detailed and accurate qualitative account of episodes of clutch performance. Studies that adopted repeat event-focused interviews with the same individual (e.g., Jackman et al., 2017), meanwhile, can provide insight into the consistent features underlying clutch performance, and how these features may develop or diminish over time.

Guiding Principles for Clutch Performance Research

Findings from the current review indicate that there are significant definitional, theoretical, and measurement issues within the field of clutch performance. These issues centre on a lack of consensus surrounding what clutch performance is, and what it is not. As a starting point in addressing these problems, we outline a number of recommendations in an effort to facilitate greater conceptual clarity. Specifically, we draw on the findings of this review to propose a number of guiding principles for future research on clutch performance.

First, clutch performance inherently requires pressure, which means that clutch performance is a psychological construct. Pressure involves the presence of situational incentives for optimal performance, and crucially, involves a subjective component (i.e., the situation is internally appraised as important; Baumeister, 1984). Accordingly, clutch performance cannot solely be measured as a behavioural outcome (such as runs scored; Deane & Palmer, 2006), as this method cannot account for subjective appraisal of situational importance. Measurement of pressure, therefore, is required when examining clutch performance, and future research should investigate if, and through what mechanisms, pressure may lead to increased performance.

Second, clutch performance is an isolated episode of performance – not an ability. Baumeister and Showers (1986) noted that 'pressure by definition focuses on a single, present performance' (p. 362). As discussed above, pressure is a requirement of clutch performance, and hence clutch performance must be an isolated episode. Further, the current review showed strong support for clutch performance as an isolated performance episode, whilst evidence for clutch performance as an ability was limited. Indeed, any examination of clutch ability inherently relies on first understanding singular episodes of clutch performance (Hibbs, 2010). Accordingly, research should examine clutch performance as an isolated performance episode, with a focus on understanding the requirements and boundaries of such an episode, before investigating the notion of clutch ability.

Third, *positive* performance is required for clutch performance. Otten (2009) defines clutch performance as 'increased or superior performance' (p. 582), whilst Hibbs (2010) specifies a 'successful performance' (p. 49). Whilst the current review demonstrated support for both of these positions (Gray et al., 2013; Swann et al., 2017b), several questions remain. For example, when considering increased or superior performance, it is unclear what magnitude performance needs to increase by, and what benchmark the performance is

compared against. It is also unclear as to what is required to constitute performance (e.g., is increased effort, or particular components of performance, sufficient?). Using 'successful performance', meanwhile, raises concerns over the extent to which clutch performance conceptually overlaps with constructs such as coping and choking-resistance (Kaiseler et al., 2009; Mesagno & Marchant, 2013). Therefore, at this stage, it is difficult to recommend the position of either Otten (2009) or Hibbs (2010). Accordingly, we recommend that researchers investigate *positive* performance under pressure. The intentions behind proposing the term *positive* are twofold. Firstly, it acts as an umbrella term that encapsulates both increased, and successful, performance. Secondly, investigating a broad range of performances is important in bringing clarity to the questions raised above. For example, one line of inquiry for future research may be examining what performance thresholds athletes and exercisers utilise to evaluate their own performance under pressure. As such, this principle is proposed with the intention to be tested, challenged and refined through future research.

Last, the role of perceived (i.e., positively appraised) performance should be considered when evaluating clutch performance. The current review included a significant body of literature that primarily reported on perceived performance (e.g., Swann et al., 2019), in addition to studies that examined objective performance (e.g., Gray et al., 2013). Indeed, neither Otten's (2009) nor Hibbs' (2010) definitions specify a distinction between perceived or objective performance. As such, it is recommended future research examines both objective and positively appraised performance. This principle should be adopted with an emphasis on understanding how athletes and exercisers judge their own performance. That is, do athletes and exercisers primarily rely on objective performance or perceived performance, or a combination of both, when evaluating their own performance under pressure.

The four guiding principles outlined above are provided as a tentative solution (Popper, 1981), and accordingly, are open to refutation. Indeed, these recommendations are

proposed with the aim to stimulate further debate around what constitutes clutch performance and help guide future research. In summary, we recommend that researchers and practitioners be critical in adopting existing definitions of clutch performance and aim to develop a refined definition and theory of clutch performance.

Strengths and Limitations

The systematic nature of the review was a strength. Efforts were taken to ensure transparency, limit author bias, and improve trustworthiness. Despite these strengths, there are also several limitations of the current review that are important to note. Firstly, this review excluded studies that were not in English or not in a peer reviewed journal, which may have created a language and publication bias. Secondly, the focus on participants in sports and exercise meant that related performance domains that may have investigated clutch performance were excluded. Third, to ensure that clutch was a primary focus of the study, the term clutch was only searched for in the title, abstract, and keyword field. Indeed, this may partly explain the relatively low return of 27 studies that were included in the present review, despite facilitation of performance under pressure being a fundamental aim of sport and exercise psychology. We recognise that studies in overlapping fields may not use the terminology of clutch performance, but rather more generic terminology (e.g., performance under pressure). However, to avoid the confounding of multiple concepts, and to limit the amount of irrelevant studies in the screening process, the focus of the present review was solely on the concept of clutch performance. Whilst the limitations of this review are recognised, at all stages steps were taken to limit these, whilst some were also inherent to the nature of the review question (e.g., a focus on sport and exercise).

Conclusion

The concept of clutch performance has experienced a substantial increase in research attention and activity over the last decade. This review demonstrated, however, that there are

804 significant definitional, conceptual, and measurement issues within the field. Specifically,
805 there appears to be a lack of clarity regarding what clutch performance is, and what clutch
806 performance is not. In response, four guiding principles were provided as a tentative solution
807 (Popper, 1981). In putting forth these principles, we seek to open debate around the concept
808 of clutch performance in an effort to move the field forward. Indeed, definitional and
809 conceptual refinement is essential to facilitate appropriate measurement of clutch
810 performance, and in turn, move the field closer to its' overarching aim: to help individuals
811 perform positively under pressure.
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828

References

Bacharach, S. B. (1989). Organizational Theories: Some Criteria for Evaluation. *Academy of Management Review*, *14*(4), 496–515. https://doi.org/10.2307/258555

Baumeister, R. F. (1984). Choking under pressure: Self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of Personality and Social Psychology*, *46*(3), 610–620. https://doi.org/10.1037/0022-3514.46.3.610

Baumeister, R. F., & Showers, C. J. (1986). A review of paradoxical performance effects: Choking under pressure in sports and mental tests. *European Journal of Social Psychology*, *16*, 361–383. https://doi.org/10.1177/1368431009355866

Beilock, S. L., & Carr, T. H. (2001). On the fragility of skilled performance: What governs choking under pressure? *Journal of Experimental Psychology: General*, *130*(4), 701–725. https://doi.org/10.1037/0096-3445.130.4.701

Birnbaum, P. (2008). Clutch Hitting and the Cramer Test. *The Baseball Resarch Journal*, *37*, 71–75.

Birnbaum, P. (2009). Players Being "Clutch" When Targeting 20 Wins. *The Baseball Research Journal*, *38*, 44–48.

Brewer, B. W., Van Raalte, J. L., Linder, D. E., & Van Raalte, N. S. (1991). Peak Performance and the Perils of Retrospective Introspection. *Journal of Sport and Exercise Psychology*, *13*(3), 227–238. https://doi.org/10.1123/jsep.13.3.227

Brooks, H. (1989). The Statistical Mirage of Clutch Hitting. *The Baseball Research Journal*, *18*. Retrieved from http://research.sabr.org/journals/the-statistical-mirage-of-clutch-hitting

Burton, D., & Naylor, S. (1997). Is anxiety really facilitative? Reaction to the myth that cognitive anxiety always impairs sport performance. *Journal of Applied Sport Psychology*, *9*(2), 295–302. https://doi.org/10.1080/10413209708406488

Cao, Z., Price, J., & Stone, D. F. (2011). Performance under pressure in the NBA. *Journal of Sports Economics*, *12*(3), 231–252. https://doi.org/10.1177/1527002511404785

Centre for Reviews and Dissemination. (2009). *Systematic reviews: CRD's guidance for undertaking reviews in health care*.

Cooper, C. L., Dewe, P. J., & O'Driscoll, M. (2001). *Stress and work organizations: A review and critique of theory, research and applications*. Sage.

Costa, G. B., Huber, M. R., & Saccoman, J. T. (2019). *Understanding Sabermetrics: An Introduction to the Science of Baseball Statistics* (2nd ed.). McFarland & Company, Inc.

Cox, R. H., Martens, M. P., & Russell, W. D. (2003). Measuring anxiety in athletes: The Revised Competitive State Anxiety Inventory-2. *Journal of Sport & Exercise Psychology*, *25*, 519–533.

Cramer, R. D. (1977). Do clutch hitters exist? *Baseball Research Journal*, *6*, 74–79.

Cramer, R. D., & Palmer, P. (2008). Clutch Hitting Revisited. *The Baseball Resarch Journal*, *37*, 85–88.

Csikszentmihalyi, M. (2002). *Flow: The psychology of optimal experience* (2nd ed.). Harper & Row.

Cunningham, G. B. (2013). Theory and theory development in sport management. *Sport Management Review*, *16*(1), 1–4. https://doi.org/10.1016/j.smr.2012.01.006

Deane, B., & Palmer, P. (2006). Still Searching for Clutch Pitchers. *The Baseball Resarch Journal*, *35*, 124–125.

Doherty, A. (2013). Investing in sport management: The value of good theory. *Sport Management Review*, *16*(1), 5–11. https://doi.org/10.1016/j.smr.2011.12.006

Dohme, L. C., Backhouse, S., Piggott, D., & Morgan, G. (2017). Categorising and defining popular psychological terms used within the youth athlete talent development literature: A systematic review. *International Review of Sport and Exercise Psychology*, *10*(1),

134–163. https://doi.org/10.1080/1750984X.2016.1185451

Drisko, J. W. (2019). Qualitative research synthesis: An appreciative and critical introduction. *Qualitative Social Work*, *0*(0), 1-18. https://doi.org/10.1177/1473325019848808

Gray, R., & Cañal-Bruland, R. (2015). Attentional focus, perceived target size, and movement kinematics under performance pressure. *Psychonomic Bulletin and Review*, *22*(6), 1692–1700. https://doi.org/10.3758/s13423-015-0838-z

Gucciardi, D. F., & Dimmock, J. A. (2008). Choking under pressure in sensorimotor skills: Conscious processing or depleted attentional resources? *Psychology of Sport and Exercise*, *9*(1), 45–59. https://doi.org/10.1016/j.psychsport.2006.10.007

Hart, S. (2017). *Celebrate Sergio: Soccer's most clutch players*. Soccer.Com. https://www.soccer.com/guide/celebrate-sergio-soccers-clutch-players

Hibbs, D. (2010). A Conceptual Analysis of Clutch Performances in Competitive Sports. *Journal of the Philosophy of Sport*, *37*(1), 47–59. https://doi.org/10.1080/00948705.2010.9714765

Higgins, J. P. T., López-López, J. A., Becker, B. J., Davies, S. R., Dawson, S., Grimshaw, J. M., McGuinness, L. A., Moore, T. H. M., Rehfuess, E. A., Thomas, J., & Caldwell, D. M. (2019). Synthesising quantitative evidence in systematic reviews of complex health interventions. *BMJ Global Health*, *4*, e000858. https://doi.org/10.1136/bmjgh-2018-000858

Hill, D. M., Carvell, S., Matthews, N., Weston, N. J. V., & Thelwell, R. R. C. (2017). Exploring choking experiences in elite sport: The role of self-presentation. *Psychology of Sport and Exercise*, *33*, 141–149. https://doi.org/10.1016/j.psychsport.2017.09.001

Hill, D. M., & Hemmings, B. (2015). A phenomenological exploration of coping responses associated with choking in sport. *Qualitative Research in Sport, Exercise and Health*,

7(4), 521–538. https://doi.org/10.1080/2159676X.2014.981573

Hubby, B., & Williamson, W. (1988). Independent studies of the reliability and validity of responses to the Professional DynaMetric Programs ProScan Survey. *Research Monograph, 9*.

Hurley, M. (2019). *There's Now More Than An Entire Game's Worth Of Tom Brady Clutch Drives In Super Bowl And Postseason*. CBS Boston. https://boston.cbslocal.com/2019/02/15/more-than-entire-game-worth-of-tom-brady-clutch-drives-in-super-bowl-and-postseason/

Jaarsma, E. A., & Smith, B. (2018). Promoting physical activity for disabled people who are ready to become physically active: A systematic review. *Psychology of Sport and Exercise, 37*, 205–223. https://doi.org/10.1016/j.psychsport.2017.08.010

Jackman, P. C., Crust, L., & Swann, C. (in press). The role of mental toughness in the occurrence of flow and clutch states in sport. *International Journal of Sport Psychology*.

Jackman, P. C., Crust, L., & Swann, C. (2017). Systematically comparing methods used to study flow in sport: A longitudinal multiple-case study. *Psychology of Sport & Exercise, 32*, 113–123. https://doi.org/10.1016/j.psychsport.2017.06.009

Jackson, R. C. (2013). Babies and bathwater: Commentary on Mesagno and Hill's proposed re-definition of "choking." *International Journal of Sport Psychology, 44*, 281–284.

Jetter, M., & Walker, J. K. (2015). Game, set, and match: Do women and men perform differently in competitive situations? *Journal of Economic Behavior and Organization, 119*, 96–108. https://doi.org/10.1016/j.jebo.2015.07.017

Kaiseler, M., Polman, R., & Nicholls, A. (2009). Mental toughness, stress, stress appraisal, coping and coping effectiveness in sport. *Personality and Individual Differences, 47*(7), 728–733. https://doi.org/10.1016/j.paid.2009.06.012

Kent, S., Devonport, T. J., Lane, A. M., Nicholls, W., & Friesen, A. P. (2018). The effects of

coping interventions on ability to perform under pressure. *Journal of Sports Science and Medicine*, *17*(1), 40–55.

Kmet, L., Lee, R., & Cook, L. (2004). *Standard Quality Assessment Criteria for evaluating primary research papers from a variety of fields*. https://doi.org/10.7939/R37M04F16

Low, W. R., Sandercock, G. R. H., Freeman, P., Winter, M. E., Butt, J., & Maynard, I. (2020). Pressure training for performance domains: A meta-analysis. *Sport, Exercise, and Performance Psychology*. https://doi.org/10.1037/spy0000202

Maher, R., Marchant, D., Morris, T., & Fazel, F. (2018). Managing pressure at the free-throw line: Perceptions of elite basketball players. *International Journal of Sport and Exercise Psychology*, *818*, 1–17. https://doi.org/10.1080/1612197X.2018.1536159

Marchant, D., Maher, R., & Wang, J. (2014). Perspectives on choking in sport. In A. G. Papaioannou & D. Hackfort (Eds.), *Routledge companion to sport and exercise psychology: Global perspectives and fundamental concepts* (pp. 446–459). Routledge.

McEwan, D., Schmaltz, R., & Ginis, K. A. M. (2012). Warming up with Pressure Improves Subsequent Clutch Performance on a Golf-Putting Task. *Advances in Physical Education*, *2*(4), 144–147. https://doi.org/10.4236/ape.2012.24025

Mesagno, C., & Hill, D. M. (2013). Definition of choking in sport: Re-conceptualization and debate. *International Journal of Sport and Exercise Psychology*, *44*, 267–277.

Mesagno, C., & Marchant, D. (2013). Characteristics of Polar Opposites: An Exploratory Investigation of Choking-Resistant and Choking-Susceptible Athletes. *Journal of Applied Sport Psychology*, *25*(1), 72–91. https://doi.org/10.1080/10413200.2012.664605

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses : The PRISMA Statement. *PLoS Medicine*, *6*(7). https://doi.org/10.1371/journal.pmed.1000097

Otten, M. (2009). Choking vs. clutch performance: a study of sport performance under

954 pressure. *Journal of Sport & Exercise Psychology*, *31*(5), 583–601.
955 Otten, M. P. (2013). Clutch Performance in Sport: A positive psychology perspective.
956 *International Journal of Sport Psychology*, *44*, 285–287.
957 Otten, M. P., & Barrett, M. E. (2013). Pitching and clutch hitting in Major League Baseball:
958 What 109 years of statistics reveal. *Psychology of Sport and Exercise*, *14*(4), 531–537.
959 https://doi.org/10.1016/j.psychsport.2013.03.003
960 Oudejans, R. R. D., Kuijpers, W., Kooijman, C. C., & Bakker, F. C. (2011). Thoughts and
961 attention of athletes under pressure: Skill-focus or performance worries? *Anxiety, Stress*
962 *and Coping*, *24*(1), 59–73. https://doi.org/10.1080/10615806.2010.481331
963 Owens, L., Stewart, C., & Huebner, E. (2016). The clutch athlete, choking and personality.
964 *Sports Coaching Review*, *6*(1), 20–35. https://doi.org/10.1080/21640629.2016.1180832
965 Perry, J. (2019). *Performing Under Pressure: Psychological Strategies for Sporting Success*
966 (1st ed.). Routledge.
967 Pluye, P., Robert, E., Cargo, M., Bartlett, G., O'Cathain, A., Griffiths, F., Boardman, F.,
968 Gagnon, M. P., & Rousseau, M. C. (2011). *Mixed Methods Appraisal Tool (MMAT):*
969 *For dissemination, application, and feedback* (Issue Part I).
970 http://mixedmethodsappraisaltoolpublic.pbworks.com/w/file/fetch/84371689/MMAT
971 2011 criteria and tutorial 2011-06-29updated2014.08.21.pdf
972 Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., & Rodgers, M. (2006). *Guidance*
973 *on the conduct of narrative synthesis in systematic reviews: A product from the ESRC*
974 *methods programme* (Issue April 2006).
975 Popper, K. R. (1981). The rationality of scientific revolutions. In I. Hacking (Ed.), *Scientific*
976 *revolutions* (pp. 80–106). Oxford University Press.
977 Professional Dynamic Programs. (2003). *Technical & research reference guide.*
978 Ruane, T. (2005). In Search of Clutch Hitting. *The Baseball Resarch Journal*, *34*, 29–36.

Safire, W. (2005). *Go To!* The New York Times. https://www.nytimes.com/2005/05/08/magazine/go-to.html

Seifreid, C., & Papatheodorou, M. (2010). The Concepts of Clutch and Choking: Recommendations for Improving Performance under Pressure. *Journal of Coaching Education*, *3*(1), 91–98.

Siddaway, A. P., Wood, A. M., & Hedges, L. V. (2019). How to Do a Systematic Review: A Best Practice Guide for Conducting and Reporting Narrative Reviews, Meta-Analyses, and Meta-Syntheses. *Annual Review of Psychology*, *70*(1), 747–770. https://doi.org/10.1146/annurev-psych-010418-102803

Sirriyeh, R., Lawton, R., Gardner, P., & Armitage, G. (2012). Reviewing studies with diverse designs: The development and evaluation of a new tool. *Journal of Evaluation in Clinical Practice*, *18*(4), 746–752. https://doi.org/10.1111/j.1365-2753.2011.01662.x

Smith, B., & McGannon, K. R. (2018). Developing rigor in qualitative research: problems and opportunities within sport and exercise psychology. *International Review of Sport and Exercise Psychology*, *11*(1), 101–121. https://doi.org/10.1080/1750984X.2017.1317357

Smith, B., & Sparkes, A. C. (2019). Interviews: qualitative interviewing in the sport and exercise sciences. In B. Smith & A. C. Sparkes (Eds.), *Routledge Handbook of Qualitative Research in Sport and Exercise* (1st ed., pp. 103–123). Routledge.

Solomonov, Y., Avugos, S., & Bar-Eli, M. (2015). Do clutch players win the game? Testing the validity of the clutch player's reputation in basketball. *Psychology of Sport and Exercise*, *16*, 130–138. https://doi.org/10.1016/j.psychsport.2014.10.004

Swann, C., Crust, L., Jackman, P., Vella, S. A., Allen, M. S., & Keegan, R. (2017a). Performing under pressure: Exploring the psychological state underlying clutch performance in sport. *Journal of Sports Sciences*, *35*(23), 2272–2280.

https://doi.org/10.1080/02640414.2016.1265661

Swann, C., Crust, L., Jackman, P., Vella, S. A., Allen, M. S., & Keegan, R. (2017b). Psychological States Underlying Excellent Performance in Sport: Toward an Integrated Model of Flow and Clutch States. *Journal of Applied Sport Psychology*, *29*(4), 375–401. https://doi.org/10.1080/10413200.2016.1272650

Swann, C., Jackman, P. C., Schweickle, M. J., & Vella, S. A. (2019). Optimal experiences in exercise: A qualitative investigation of flow and clutch states. *Psychology of Sport and Exercise*, *40*, 87–98. https://doi.org/10.1016/j.psychsport.2018.09.007

Swann, C., Keegan, R., Crust, L., & Piggott, D. (2016). Psychological states underlying excellent performance in professional golfers: "Letting it happen" vs. "making it happen." *Psychology of Sport and Exercise*, *23*, 101–113. https://doi.org/10.1016/j.psychsport.2015.10.008

Swann, C., Moran, A., & Piggott, D. (2015). Defining elite athletes: Issues in the study of expert performance in sport psychology. *Psychology of Sport and Exercise*, *16*(P1), 3–14. https://doi.org/10.1016/j.psychsport.2014.07.004

Swann, C., Piggott, D., Schweickle, M., & Vella, S. A. (2018). A Review of Scientific Progress in Flow in Sport and Exercise: Normal Science, Crisis, and a Progressive Shift. *Journal of Applied Sport Psychology*, *30*(3), 249–271. https://doi.org/10.1080/10413200.2018.1443525

Thomas, J., Harden, A., & Newman, M. (2012). Synthesis: Combining results systematically and appropriately. In D. Gough, S. Oliver, & J. Thomas (Eds.), *An introduction to systematic reviews* (1st ed., pp. 179–226). SAGE Publications.

Thomas, O., Hanton, S., & Jones, G. (2002). An alternative approach to short-form self-report assessment of competitive anxiety. *International Journal of Sport Psychology*, *33*, 325–336.

Wacker, J. G. (2004). A theory of formal conceptual definitions: Developing theory-building measurement instruments. *Journal of Operations Management*, *22*(6), 629–650. https://doi.org/10.1016/j.jom.2004.08.002

Wallace, S., Caudill, S. B., & Mixon, F. G. (2013). Homo certus in professional basketball? Empirical evidence from the 2011 NBA Playoffs. *Applied Economics Letters*, *20*(7), 642–648. https://doi.org/10.1080/13504851.2012.727965

West, J., & Libby, B. (1969). *Mr. Clutch: The Jerry West Story* (3rd ed.). Prentice-Hall.

Woodyard, E. (2018). *20 years later, Utah Jazz's 1998 NBA Finals team still haunted by Jordan's 'Last Shot.'* Desert News. https://www.deseretnews.com/article/900021505/20-years-later-jazz-s-1998-nba-finals-team-still-haunted-by-jordan-s-last-shot.html

World Health Organisation. (2018). *Global action plan on physical activity 2018-2030. More active people for a healthier world.*

Worthy, D. A., Markman, A. B., & Maddox, T. W. (2009). What is pressure? Evidence for social pressure as a type of regulatory focus. *Psychonomic Bulletin and Review*, *16*(2), 344–349. https://doi.org/10.3758/PBR.16.2.344

Yarrow, M. R., Campbell, J. D., & Burton, R. V. (1970). Recollections of Childhood a Study of the Retrospective Method. *Monographs of the Society for Research in Child Development*, *35*(5), iii. https://doi.org/10.2307/1165649